DOCUMENT RESUME

ED 197 515 EC 131 717

AUTHOR Ysseldyke, James E.: And Others
TITLE The Use of Technically Adequate Tests in Psychoeducational Decision Making.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
REPORT NO IRLD-RR-28
PUB DATE Apr 80
CONTRACT 300-77-0491
NOTE 36p.: For related documents, see EC 131 709-720.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Disabilities: *Evaluation Methods: Exceptional Child Research: Professional Personnel: Screening Tests: *Student Evaluation: *Student Placement: *Testing Problems

ABSTRACT

Results of three separate studies were analyzed to ascertain the technical adequacy of tests used by professionals to make screening, placement/classification, instructional planning, and evaluation decisions for handicapped students. In each investigation, the frequency of usage of technically adequate instruments was addressed. The finding that various professionals employ a large number of technically inadequate measures is discussed in terms of implications for current assessment and decision making practices. (Author)

University of Minnesota

Research Report No. 28

# THE USE OF TECHNICALLY ADEQUATE TESTS IN PSYCHOEDUCATIONAL DECISION MAKING

James E. Ysseldyke, Richard R. Regan, and Sherry Z. Schwartz


**Institute for Research on Learning Disabilities**

IRLD

Director: James E. Ysseldyke

Associate Director: Phyllis K. Mirkin

The Institute for Research on Learning Disabilities is supported by a contract (300-77-0491) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U.S. Office of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled children. Research activities are organized into eight major areas:

I. Adequacy of Norm-Referenced Data for Prediction of Success

II. Computer Simulation Research on the Assessment/Decision-making/Intervention Process

III. Comparative Research on Children Labeled LD and Children Failing Academically but not Labeled LD

IV. Surveys on In-the-Field Assessment, Decision Making, and Intervention

V. Ethological Research on Placement Team Decision Making

VI. Bias Following Assessment

VII. Reliability and Validity of Formative Evaluation Procedures

VIII. Data-Utilization Systems in Instructional Programming

Additional information on these research areas may be obtained by writing to the Editor at the Institute.

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Research Report No. 28

# THE USE OF TECHNICALLY ADEQUATE TESTS IN PSYCHOEDUCATIONAL DECISION MAKING

James E. Ysseldyke, Richard R. Regan, and Sherry Z. Schwartz

Institute for Research on Learning Disabilities

University of Minnesota

April, 1980

## Abstract

The results of three separate studies were analyzed to ascertain the technical adequacy of tests used by professionals. In each investigation, the frequency of usage of technically adequate instruments was addressed. The finding that various professionals employ a large number of technically inadequate measures is discussed in terms of implications for current assessment and decision-making practices.

# The Use of Technically Adequate Tests in Psychoeducational Decision Making

Within the last decade those who assess and make psychoeducational decisions about students have had to demonstrate increased accountability at the level of the individual. First, as a result of litigation, and more recently as a result of legislation, decision makers are having to lay bare their assessment and decision-making activities. Repeatedly, we observe criticism of the technical adequacy of tests used to make screening, placement/classification, instructional planning, and evaluation decisions for students (Arter & Jenkins, 1977; Salvia & Ysseldyke, 1978; Ysseldyke, 1973, 1978a, 1978b, 1978c, 1979; Ysseldyke, Algozzine, Regan, & Potter, 1979). Both the Office of Civil Rights Regulations, published to accompany section 504 of the Rehabilitation Act of 1973, and the "Protection in Evaluation Procedures Provisions" of the Education for all Handicapped Children Act of 1975 specify that tests must have been validated for the purposes for which they are used.

While there have been repeated exhortations regarding the importance of using technically adequate assessment instruments, there are few characterizations of the technical adequacy of tests used by decision makers. This investigation used multiple methodologies to ascertain the extent to which decision makers use technically adequate tests in the process of making decisions about students.

## Method

### Design

Three separate studies were conducted to ascertain the frequency of

usage of technically adequate instruments by professionals.

The first study used a self-report methodology. Subjects were asked to identify those tests used most often (1) in general, and with students who were referred for (2) academic and (3) behavior problems.

The second study used a simulated decision-making procedure. Subjects were given referral data on a hypothetical student who demonstrated either (1) academic or (2) behavior problems, and then used a computer terminal to select assessment data they wanted on a student.

The third study used a questionnaire methodology in which professionals from federally funded model programs for learning disabled students were asked to identify those tests used to assess students.

Subjects

Subjects for Study One were 65 decision makers from Virginia. A variety of disciplines was represented, including 31 school psychologists or school psychology interns, 15 regular or special education teachers, eight support personnel (nurses, counselors, social workers, administrators), and 11 individuals who did not specify their role.

Subjects for Study Two were 159 educational personnel from the greater Minneapolis/St. Paul metropolitan area, all of whom had participated in at least two placement team meetings. Occupational groups represented in the sample were administrators, regular education teachers, special education teachers, school psychologists, and support personnel.

Data for Study Three were obtained by mailing a questionnaire to 52 demonstration programs for learning disabled students. Questionnaires were received from 44 model programs in 26 states. Since six

programs reported that they did not assess students, usable data from 38 centers were considered.

Procedures

The multiple methodology approach enabled us to analyze data on the same set of questions from different samples in different ways. The subjects for Study One were given a list of 49 commonly used assessment instruments in seven domains, and then were asked three questions. First, they were asked to rank order the five devices they most frequently used in assessing elementary-age students. Next, they were asked to rank order the five devices they used to assess students referred for academic problems. Finally, they were asked to rank order the five devices they used most often to assess students referred for behavior problems.

A computer-simulated decision-making program was used in Study Two. Subjects were given data on a hypothetical referred student who evidenced either academic or behavior problems. Subjects were allowed to access both quantitative and qualitative data on student performance from the same tests that were included in Study One. After receiving information on pupil performance on as many devices as they wished, subjects were asked to make eligibility, classification, and prognostic decisions. Data were accessed via a telray remote terminal attached by phone to a Cybernet computer. The computer recorded those devices selected by the participants.

In Study Three, personnel from model programs were asked to identify those tests used for the purpose of making screening, classification, intervention planning, and evaluation decisions.

Technical Adequacy Criteria

Technical adequacy of the tests was evaluated on three dimensions: norms, reliability, and validity. Table 1 summarizes our evaluation of the technical adequacy of the tests based on the APA Standards for Educational and Psychological Tests, and criteria specified by Salvia and Ysseldyke (1978) and Thurlow and Ysseldyke (1979).

- - - - - - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - - - - - -

## Results

In all three studies, subjects were able to sample tests from each of the seven domains more than once. This flexibility limited data analysis to the calculation of descriptive statistics.

### Study One

Table 2 presents the specific devices selected most often as the first device, the second device, and so on, as a function of referral information (general, academic, or behavioral). Also included is the weighted rank of each device, derived by assigning a position rank and summing. As is evident in the table, the WISC-R was selected most often, whether for general use or for use with students with specific problems.

- - - - - - - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - - - - - - -

Each selected device was rated for technical adequacy with respect to reliability and validity. A "+" was assigned for technical adequacy in each category and a "-" for technical inadequacy. Tables 3 through 5

identify, by rank, the number of adequate, inadequate, and other (special condition and criterion referenced) devices, as well as the percentage of adequate, inadequate, and other devices selected.

- - - - - - - - - - - - - - - -

Insert Tables 3-5 about here

- - - - - - - - - - - - - - - -

The most striking result was the decreasing use of technically adequate devices over time. The first test selected was nearly always technically adequate with regard to norms, reliability, and validity. The device selected by 95 percent of the subjects had technically adequate norms, the device selected by 94 percent of the subjects was reliable, and the device selected by 94 percent of the subjects was valid. More than half of the tests selected second were reliable, but in all other instances fewer than half the tests selected were technically adequate, with regard to norms, reliability, or validity.

Study Two

Participants in this investigation were provided data on a hypothetical referred student who evidenced either academic or behavior problems. Each subject was provided with a list of tests (see Table 1) and was allowed to access both quantitative and qualitative information on student performance on tests of their choice. Participants were permitted to select as many devices as they wished. Data collected during the investigation addressed the frequency of specific test use (1) with students referred for academic problems, and (2) with students referred for behavioral problems. Ranks were assigned to tests in a weighted

manner. The tests selected most often and their weighted scales are listed in Table 6. As in Study One, the WISC-R was selected most often, regardless of the student's problems.

- - - - - - - - - - - - - - - - - -

Insert Table 6 about here

- - - - - - - - - - - - - - - - - -

Each device selected was rated on technical adequacy for norms, reliability, and validity using the criteria stipulated in Study One. Tables 7 through 9 identify, in order of selection, the number and percentage of adequate, inadequate, and other (special condition and criterion referenced) devices selected in Study Two.

- - - - - - - - - - - - - - - - - -

Insert Tables 7-9 about here

- - - - - - - - - - - - - - - - - -

These results reflect a pattern of test usage comparable to that identified in Study One. Devices initially selected were adequate with regard to norms, validity, and reliability for either the academic or behavioral case. However, as more devices were reviewed (i.e., fourth, fifth, or sixth selection), there was a marked decline in the number of devices that were technically adequate on the three dimensions under consideration; correspondingly, the number of technically inadequate devices increased.

Although subjects in both academic and behavioral conditions used a greater frequency of technically adequate than inadequate devices early in the data collection and review process, there was a notable difference in the relative frequencies of technically adequate measures

reviewed for the academic and behavioral cases. The difference between the two conditions may be accounted for by the large number of "other" devices (i.e., special condition, criterion referenced) reviewed by subjects in the behavioral condition.

The analysis of the results indicated that regardless of condition, subjects selected technically adequate devices most frequently early in the review process and increased their use of technically inadequate measures as the decision-making process continued. Similarly, decisions for children with behavioral problems tended to be based on technically inadequate measures more often than decisions for children with academic problems.

Study Three

The assessment data used by CSDCs and the decisions to which those data were applied are listed in Table 10. A review of the data indicated that specific assessment devices and/or strategies were used for all types of decisions, ranging from screening to program evaluation (Thurlow & Ysseldyke, 1979). Thurlow and Ysseldyke found that norm-referenced tests were among the two most frequently used sources of data in all decision areas, except instructional program decisions, where criterion-referenced tests, informal devices, and observations were used more often.

- - - - - - - - - - - - - - - - - -

Insert Table 10 about here

- - - - - - - - - - - - - - - - - -

Assessment devices reported by five or more CSDCs were evaluated in terms of their technical adequacy on three dimensions: norms, reliability, and validity. Technical characteristics of the various assessment

devices identified were evaluated in accordance with the criteria specified by Salvia and Ysseldyke (1978), Ysseldyke (1978a), and the APA (1972) Standards. Evaluation of the 18 specific instruments used by five or more centers indicated only five (26.3%) had technically adequate norms, six (31.5) had reliability adequate for use in decision making, and five (26.3%) had technically adequate validity. Of the four devices used by at least half of the CSDCs (Key Math, PIAT, WISC-R, and WRAT), two had technically adequate norms, three had adequate reliability, and two had adequate validity (Thurlow & Ysseldyke, 1979).

Of the seven most frequently used devices identified by the CSDCs, regardless of the decisions for which they were employed, three had adequate norms, five had acceptable reliability, and four demonstrated adequate validity. An interesting result noted in this investigation and the other two studies reported here is that the WISC-R appears to be among the most frequently used measures regardless of the decision to be made.

## Discussion

The results of this comparative evaluation of assessment practices of various professionals leave a number of issues related to assessment unresolved and the appropriateness of current psychoeducational decision-making practices in doubt. Presently, no controls exist for monitoring the publication of tests with inadequate norms, reliability, and/or validity. Salvia and Ysseldyke (1978) pointed out that a number of the currently popular assessment devices used by educators are technically inadequate based on professional standards for best practices (APA, 1972). It seems an obvious requirement and recommended practice that professionals who engage in assessment of children should use technically adequate devices.

However, results derived from this inquiry suggest that professionals rarely attend to or consider the technical merits of assessment devices for the purpose of decision making.

Studies One and Two provided evidence of the decline in use of technically adequate measures after the first or second selection. Of those technically adequate measures selected by most professionals early in the assessment process, the Wechsler Intelligence Scale for Children - Revised accounted for a significant portion of those assessment instruments deemed technically adequate. The disproportionate use of this technically adequate measure masks the overall magnitude of use of technically inadequate devices. Professionals clearly employ a large number of technically inadequate measures; of the limited number of technically adequate measures available, a few appear to be used extensively.

The burden of appropriate selection and use of assessment measures clearly rests with the professional who engages in psychoeducational assessment. The results reported here suggest that current psychoeducational assessment and decision-making practices lack the technical rigor critical to the process. In addition, this analysis highlights the diversity of assessment strategies professionals employ when addressing the same referral problem.

Participants in these studies were all individuals who had already participated in making placement decisions. We believe it is imperative that increasing attention be given in both inservice and preservice training to the importance of technical adequacy in selection of instruments for use in decision making. Technical adequacy is but one aspect

of the psychoeducational assessment and decision-making process. In light of current litigation and legislative mandates, comprehensive education in all aspects of assessment and decision making is important.

## References

APA. Standards for educational and psychological tests and manuals. Washington, D.C.: APA, 1972.

Arter, J. A., & Jenkins, J. R. Examining the benefits and prevalence of modality considerations in special education. Journal of Special Education, 1977, 11, 281-298.

Salvia, J., & Ysseldyke, J. Assessment in special and remedial education. Boston: Houghton-Mifflin, 1978.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for learning disabled students. Learning Disability Quarterly, 1979, 2, 15-24.

Ysseldyke, J. E. Farewell to the psychometric robot: Training implications for school psychologists. Pennsylvania Psychologist, 1973, 10-13.

Ysseldyke, J. E. Implementation of the nondiscriminatory assessment provisions of Public Law 94-142. In Developing criteria for the evaluation of protection in evaluation procedures provisions. Washington, D.C.: Department of Health, Education, and Welfare, United States Office of Education, Bureau of Education for the Handicapped, 1978. (a)

Ysseldyke, J. E. Remediation of ability deficits in learning disabled adolescents: Some major questions. In L. Mann, L. Goodman, & J. L. Wiederholt (Eds.), The learning disabled adolescent. Boston: Houghton-Mifflin, 1978. (b)

Ysseldyke, J. E. Assessment of retardation. In J. Neisworth & R. M. Smith (Eds.), Retardation: Issues, assessment, and intervention. New York: McGraw-Hill, 1978. (c)

Ysseldyke, J. E. Issues in psychoeducational assessment. In D. Reschly & G. Phye (Eds.), School psychology: Perspectives and issues. New York: Academic Press, 1979.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, P. Technical adequacy of tests used in simulated decision making (Research Report No. 9). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Table 1

Technical Adequacy of Devices

| Test | Norms | Reliability | Validity |
|---|---|---|---|
| **Intelligence Tests** | | | |
| Stanford Binet | + | - | - |
| WISC-R | + | + | + |
| Slosson | - | - | - |
| McCarthy Scales of Children's Abilities | + | + | + |
| Full Range Picture Vocabulary Test | - | - | - |
| Quick Test | - | - | - |
| Peabody Picture Vocabulary Test | - | + | + |
| Goodenough-Harris Drawing Test | - | - | - |
| Henmon-Nelson Tests of Mental Ability | - | - | - |
| Kuhlmann-Anderson Intelligence Tests | + | + | + |
| Otis-Lennon Mental Ability Test | + | + | + |
| Primary Mental Abilities Test | - | + | + |
| **Achievement Tests** | | | |
| California Achievement Test | - | + | - |
| Iowa Test of Basic Skills | + | - | - |
| Metropolitan Achievement Test | - | + | - |
| Stanford Achievement Test | + | + | + |
| Gates-MacGinitie Reading Tests | - | + | - |
| Peabody Individual Achievement Tests | + | + | + |
| Wide Range Achievement Test | - | + | - |
| Gray Oral Reading Test | - | - | - |
| Gilmore Oral Reading Test | - | - | - |
| Gates-McKillop Reading Diagnostic Tests | - | - | - |
| Durrell Analyses of Reading Difficulty | - | - | - |
| Stanford Diagnostic Reading Test | + | + | + |
| Diagnostic Reading Scales | - | - | - |
| Woodcock Reading Mastery Test | + | + | + |
| Key Math Diagnostic Arithmetic Test | - | - | - |
| Stanford Diagnostic Mathematics Test | + | + | + |
| Diagnosis: An Instructional Aid in Math | CR | CR | CR |
| **Perceptual-Motor Tests** | | | |
| Bender Visual-Motor Gestalt | - | - | - |
| Developmental Test of Visual Perception | - | - | - |
| Memory for Designs Test | - | - | - |
| Developmental Test of Visual-Motor Integration | - | - | - |
| Purdue Perceptual-Motor Survey | - | - | - |

| Test | Norms | Reliability | Validity |
|---|---|---|---|
| **Behavioral Recordings** | | | |
| Frequency Counting or Event Recordings | SC | SC | SC |
| Interval or Time Samplings | SC | SC | SC |
| Permanent Products | SC | SC | SC |
| Peterson-Quay Behavior Problem Checklist | - | - | - |
| **Personality Tests** | | | |
| Piers-Harris Self-Concept Scale | - | - | - |
| Rorschach-Inkblot Technique | - | - | - |
| School Apperception Method | - | - | - |
| Thematic Apperception Test | - | - | - |
| **Adaptive Behavior Scales** | | | |
| AAMD Adaptive Behavior Scale | - | - | - |
| AAMD Adaptive Behavior Scale (School Version) | + | - | - |
| Vineland Social Maturity Scale | - | - | - |
| **Language Tests** | | | |
| Goldman-Fristoe Test of Articulation | CR | + | + |
| Auditory Discrimination Test | - | - | - |
| Northwestern Syntax Screening Test | - | - | - |
| Illinois Test of Psycholinguistic Abilities | - | - | - |

Note: + = technically adequate
- = technically inadequate
CR = criterion referenced
SC = special case

Table 2

Frequency of Specific Test Usage as a Function of Referral Information

| General | Academic | Behavioral |
|---|---|---|
| (1) Wechsler Intelligence Scale for Children - Revised (302) | Wechsler Intelligence Scale for Children - Revised (251) | Wechsler Intelligence Scale for Children - Revised (179) |
| (2) Bender Visual-Motor Gestalt Test (109) | Peabody Individual Achievement Tests (82) | Frequency Counts or Event Recordings (141) |
| (3) Wide Range Achievement Test (100) | Key Math Diagnostic Arithmetic Test (58) | Interval or Time Samplings (72) |
| (4) Peabody Individual Achievement Tests (91) | Woodcock Reading Mastery Test (67) | AAMD Adaptive Behavior Scale (School Version) (55) |
| (5) Stanford-Binet Intelligence Scale (68) | Developmental Test of Visual-Motor Integration (51) | Piers-Harris Self-Concept Scale (49) |
| (6) Illinois Test of Psycholinguistic Abilities (39) | Wide Range Achievement Test (47) | Peterson-Quay Behavior Problem Checklist (48) |
| (7) Peabody Picture Vocabulary Test (30) | Bender Visual-Motor Gestalt Test (41) | Bender Visual-Motor Gestalt Test (44) |

Note: Values in parentheses represent the weighted rankings of individual devices.

Order of Frequency of Selection

Table 3

Frequency of Use of Devices According to Technical Adequacy of Norms

| | Adequate | | | Inadequate | | | Other | | |
|---|---|---|---|---|---|---|---|---|---|
| | General | Academic | Behavioral | General | Academic | Behavioral | General | Academic | Behavioral |
| (1) | 62 (.95) | 53 (.84) | 33 (.54) | 3 (.05) | 10 (.16) | 14 (.23) | 0 (.00) | 0 (.00) | 14 (.23) |
| (2) | 29 (.45) | 33 (.53) | 11 (.18) | 34 (.52) | 30 (.48) | 28 (.46) | 2 (.03) | 0 (.00) | 22 (.36) |
| (3) | 15 (.24) | 16 (.26) | 16 (.26) | 47 (.73) | 46 (.74) | 33 (.54) | 2 (.03) | 0 (.00) | 12 (.20) |
| (4) | 15 (.23) | 18 (.30) | 10 (.17) | 44 (.69) | 40 (.65) | 39 (.66) | 5 (.08) | 3 (.05) | 10 (.17) |
| (5) | 10 (.16) | 11 (.20) | 18 (.37) | 44 (.72) | 41 (.75) | 29 (.59) | 7 (.12) | 3 (.05) | 2 (.04) |
| | | 12 (.24)* | | | 32 (.66)* | | | 5 (.10)* | |

Selections

* These figures represent the number of the 49 devices available during the investigation and their technical characteristics relative to norms. Numbers in parentheses indicate percent of the total available.

Table 4

Frequency of Use of Devices According to Technical Adequacy of Reliability

| Selection | Adequate | | | Inadequate | | | Other | | |
|---|---|---|---|---|---|---|---|---|---|
| | General | Academic | Behavioral | General | Academic | Behavioral | General | Academic | Behavioral |
| (1) | 61 (.94) | 53 (.84) | 33 (.54) | 4 (.06) | 10 (.16) | 14 (.23) | 0 (.00) | 0 (.00) | 14 (.23) |
| (2) | 38 (.59) | 42 (.67) | 11 (.18) | 25 (.38) | 21 (.33) | 28 (.46) | 2 (.03) | 0 (.00) | 22 (.36) |
| (3) | 22 (.34) | 17 (.27) | 13 (.21) | 40 (.63) | 45 (.73) | 36 (.59) | 2 (.03) | 0 (.00) | 12 (.20) |
| (4) | 23 (.36) | 22 (.36) | 12 (.20) | 38 (.59) | 37 (.61) | 37 (.63) | 3 (.05) | 2 (.03) | 10 (.17) |
| (5) | 27 (.44) | 19 (.35) | 17 (.35) | 32 (.53) | 34 (.62) | 30 (.61) | 2 (.03) | 2 (.03) | 2 (.04) |
| | | 16 (.33)* | | | 29 (.59)* | | | 4 (.08)* | |

* These figures represent the number of the 49 devices available during the investigation and their technical characteristics relative to reliability. Numbers in parentheses indicate percent of the total available.

Table 5

Frequency of Use of Devices According to Technical Adequacy of Validity

| Adequate | | | Inadequate | | | Other | | |
|---|---|---|---|---|---|---|---|---|
| General | Academic | Behavioral | General | Academic | Behavioral | General | Academic | Behavioral |
| 61 (.94) | 59 (.79) | 30 (.49) | 4 (.06) | 13 (.21) | 17 (.28) | 0 (.00) | 0 (.00) | 14 (.23) |
| 23 (.35) | 30 (.48) | 6 (.10) | 40 (.62) | 33 (.52) | 33 (.54) | 2 (.03) | 0 (.00) | 22 (.36) |
| 13 (.20) | 14 (.23) | 9 (.15) | 49 (.77) | 48 (.77) | 40 (.65) | 2 (.03) | 0 (.00) | 12 (.20) |
| 10 (.15) | 17 (.28) | 9 (.15) | 51 (.80) | 42 (.69) | 40 (.68) | 3 (.05) | 2 (.03) | 10 (.17) |
| 14 (.23) | 14 (.25) | 14 (.29) | 45 (.74) | 39 (.71) | 33 (.67) | 2 (.03) | 2 (.04) | 2 (.04) |
| | 12 (.24)* | | | 33 (.67)* | | | 4 (.08)* | |

figures represent the number of the 49 devices available during the investigation and their technical
cteristics relative to validity. Numbers in parentheses indicate percent of the total available.

Table 6

Frequency of Specific Test Usage as a Function of Referral Information

| Order of Frequency of Selection | Academic | Behavioral |
|---|---|---|
| (1) | Wechsler Intelligence Scale for Children-Revised (268) | Wechsler Intelligence Scale for Children-Revised (190) |
| (2) | Stanford-Binet Intelligence Scale (96) | Stanford-Binet Intelligence Scale (100) |
| (3) | Bender Visual-Motor Gestalt Test (77) | Frequency Counting or Event Recordings (95) |
| (4) | Peabody Individual Achievement Tests (63) | Wide Range Achievement Test (58) |
| (5) | Wide Range Achievement Test (60) | Peterson-Quay Behavior Problem Checklist (57) |
| (6) | Iowa Test of Basic Skills (45) | Bender Visual-Motor Gestalt Test (55) |
| (7) | Key Math Diagnostic Arithmetic Test (39) | Piers-Harris Self-Concept Scale (50) |

Note: Values in parentheses represent the weighted rankings of individual devices.

Table 7

Frequency of Use of Devices According to

Technical Adequacy of Norms

| Order of Selection | Adequate Academic | | Adequate Behavioral | | Inadequate Academic | | Inadequate Behavioral | | Other Academic | | Other Behavioral | | Total Devices Selected |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | 72 | (.90) | 50 | (.63) | 8 | (.10) | 14 | (.18) | 0 | (.00) | 15 | (.19) | 159 |
| (2) | 36 | (.47) | 36 | (.46) | 39 | (.51) | 36 | (.46) | 2 | (.02) | 6 | (.08) | 155 |
| (3) | 23 | (.30) | 27 | (.36) | 51 | (.67) | 39 | (.53) | 3 | (.03) | 8 | (.11) | 151 |
| (4) | 14 | (.19) | 11 | (.15) | 52 | (.72) | 54 | (.75) | 6 | (.09) | 7 | (.11) | 144 |
| (5) | 7 | (.11) | 14 | (.22) | 54 | (.84) | 44 | (.68) | 3 | (.05) | 7 | (.11) | 129 |
| (6) | 8 | (.16) | 11 | (.19) | 40 | (.78) | 40 | (.68) | 3 | (.06) | 8 | (.13) | 110 |
| (7) | 13 | (.36) | 5 | (.12) | 21 | (.58) | 31 | (.74) | 2 | (.06) | 6 | (.14) | 78 |
| (8) | 2 | (.09) | 5 | (.18) | 16 | (.73) | 19 | (.70) | 4 | (.18) | 3 | (.11) | 49 |
| (9) | 3 | (.27) | 2 | (.17) | 8 | (.73) | 9 | (.75) | 0 | (.00) | 1 | (.08) | 23 |
| (10) | 1 | (.25) | 0 | (.00) | 3 | (.75) | 5 | (1.00) | 0 | (.00) | 0 | (.00) | 9 |
| (11) | 1 | (.50) | 0 | (.00) | 1 | (.50) | 1 | (1.00) | 0 | (.00) | 0 | (.00) | 3 |
| | 12 (.24)* | | | | 32 (.66)* | | | | 5 (.10)* | | | | |

*These figures represent the number of the 49 devices available during the simulated diagnostic session and their technical characteristics relative to norms. Numbers in parentheses indicate percent of the total available.

Table 8

Frequency of Use of Devices According to Technical Adequacy of Reliability

| Order of Selection | Adequate Academic | | Adequate Behavioral | | Inadequate Academic | | Inadequate Behavioral | | Other Academic | | Other Behavioral | | Total Devices Selected |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | 54 | (.68) | 38 | (.48) | 26 | (.32) | 26 | (.33) | 0 | (.00) | 15 | (.19) | 159 |
| (2) | 46 | (.61) | 43 | (.55) | 29 | (.36) | 29 | (.37) | 2 | (.03) | 6 | (.08) | 155 |
| (3) | 31 | (.40) | 30 | (.40) | 43 | (.56) | 36 | (.49) | 3 | (.04) | 8 | (.11) | 151 |
| (4) | 17 | (.24) | 16 | (.22) | 50 | (.69) | 49 | (.68) | 5 | (.07) | 7 | (.10) | 144 |
| (5) | 8 | (.12) | 9 | (.14) | 53 | (.83) | 49 | (.75) | 3 | (.05) | 7 | (.11) | 129 |
| (6) | 3 | (.06) | 11 | (.19) | 45 | (.86) | 41 | (.70) | 3 | (.08) | 7 | (.12) | 110 |
| (7) | 9 | (.25) | 2 | (.05) | 25 | (.69) | 34 | (.81) | 2 | (.06) | 6 | (.14) | 78 |
| (8) | 3 | (.14) | 3 | (.11) | 15 | (.68) | 22 | (.82) | 4 | (.18) | 2 | (.07) | 49 |
| (9) | 1 | (.09) | 2 | (.17) | 10 | (.91) | 9 | (.75) | 0 | (.00) | 1 | (.08) | 23 |
| (10) | 0 | (.00) | 0 | (.00) | 4 | (1.00) | 5 | (1.00) | 0 | (.00) | 0 | (.00) | 9 |
| (11) | 1 | (.50) | 0 | (.00) | 1 | (.50) | 1 | (1.00) | 0 | (.00) | 0 | (.00) | 3 |
| | 16 (.33)* | | | | 29 (.59)* | | | | 4 (.08)* | | | | |

*These figures represent the number of the 49 devices available during the simulated diagnostic session and their technical characteristics relative to reliability. Number in parentheses indicate percent of the total available.

Table 9

Frequency of Use of Devices According to

Technical Adequacy of Validity

| Order of Selection | Adequate | | | | Inadequate | | | | Other | | | | Total Devices Selected |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Academic | | Behavioral | | Academic | | Behavioral | | Academic | | Behavioral | | |
| (1) | 52 | (.65) | 35 | (.44) | 28 | (.35) | 29 | (.37) | 0 | (.00) | 15 | (.19) | 159 |
| (2) | 32 | (.42) | 28 | (.36) | 43 | (.56) | 44 | (.56) | 2 | (.02) | 6 | (.08) | 155 |
| (3) | 20 | (.26) | 19 | (.26) | 54 | (.70) | 47 | (.64) | 3 | (.04) | 8 | (.11) | 151 |
| (4) | 12 | (.17) | 10 | (.14) | 55 | (.76) | 55 | (.76) | 5 | (.07) | 7 | (.10) | 144 |
| (5) | 5 | (.08) | 7 | (.11) | 56 | (.88) | 51 | (.78) | 3 | (.04) | 7 | (.11) | 129 |
| (6) | 2 | (.04) | 9 | (.15) | 46 | (.90) | 43 | (.73) | 3 | (.06) | 7 | (.12) | 110 |
| (7) | 8 | (.22) | 2 | (.05) | 26 | (.72) | 34 | (.81) | 2 | (.06) | 6 | (.14) | 78 |
| (8) | 3 | (.14) | 3 | (.11) | 15 | (.68) | 22 | (.82) | 4 | (.18) | 2 | (.07) | 49 |
| (9) | 1 | (.09) | 2 | (.17) | 10 | (.91) | 9 | (.75) | 0 | (.00) | 1 | (.08) | 23 |
| (10) | 0 | (.00) | 0 | (.00) | 4 | (1.00) | 5 | (1.00) | 0 | (.00) | 0 | (.00) | 9 |
| (11) | 1 | (.50) | 0 | (.00) | 1 | (.50) | 1 | (1.00 | 0 | (.00) | 0 | (.00) | 3 |
| | 12 (.24)* | | | | 33 (.67)* | | | | 4 (.08)* | | | | |

*These figures represent the number of the 49 devices available during the simulated diagnostic session and their technical characteristics relative to validity. Number in parentheses indicate percent of the total available.

Table 10

Percentages of Different Assessment Instruments used in Decision Making[a]

| Instrument | CSDCs[b] Using | Decision for Which Used by CSDCs[c] | | | | |
|---|---|---|---|---|---|---|
| | | Scrng | Placmt | Instruc Prog | Pupil Eval | Prog Eval |
| WISC/WISC-R | 64 | 44 | 80 | 48 | 56 | 8 |
| Key Math | 59 | 30 | 56 | 78 | 70 | 35 |
| WRAT | 59 | 48 | 60 | 39 | 56 | 39 |
| Informal | 59 | 61 | 65 | 91 | 87 | 56 |
| PIAT | 54 | 52 | 71 | 36 | 76 | 48 |
| Woodcock Reading | 38 | 40 | 80 | 67 | 60 | 40 |
| PPVT | 33 | 77 | 38 | 38 | 46 | 8 |
| Beery | 26 | 50 | 60 | 40 | 40 | 10 |
| Wepman | 23 | 56 | 89 | 67 | 78 | 11 |
| Brigance | 20 | 38 | 75 | 100 | 62 | 0 |
| Detroit | 20 | 38 | 75 | 75 | 62 | 25 |
| ITPA | 20 | 25 | 88 | 75 | 75 | 12 |
| WAIS | 15 | 50 | 67 | 33 | 67 | 0 |
| Slosson | 15 | 50 | 67 | 50 | 50 | 50 |
| Piers-Harris | 20 | 25 | 50 | 50 | 38 | 38 |
| Bender | 13 | 40 | 80 | 80 | 60 | 20 |
| Carrow | 13 | 60 | 80 | 100 | 100 | 20 |
| Spache | 13 | 60 | 80 | 80 | 80 | 20 |
| Stanford-Binet | 13 | 40 | 80 | 60 | 60 | 0 |

[a] Table includes only those instruments mentioned by five or more CSDCs.

[b] Percentages reflect numbers of CSDCs listing each instrument.

[c] Percentages reflect numbers of CSDCs using instrument for each decision based only on those listing the instrument.

# PUBLICATIONS

## Institute for Research on Learning Disabilities
## University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor
IRLD
350 Elliott Hall
75 East River Road
University of Minnesota
Minneapolis, Minnesota 55455

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota roundtable conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Note: Monographs No. 1 - 6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.

Sutherland, J. H., Algozzine, B., Ysseldyke, J. E., & Young, S. What can I say after I say LD? (Research Report No. 16). December, 1979.

Deno, S. L., & Mirkin, P. K. Data-based IEP development: An approach to substantive compliance (Monograph No. 13). December, 1979.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. The influence of test scores and naturally-occurring pupil characteristics on psycho-educational decision making with children (Research Report No. 17). December, 1979.

Algozzine, B., & Ysseldyke, J. E. Decision makers' prediction of students' academic difficulties as a function of referral information (Research Report No. 18). December, 1979.

Ysseldyke, J. E., & Algozzine, B. Diagnostic classification decisions as a function of referral information (Research Report No. 19). January, 1980.

Deno, S. L., Chiang, B., Mirkin, P. K., & Lowry, L. Relationships among simple measures of reading and performance on standardized achievement tests (Research Report No. 20). January, 1980.

Deno, S. L., Lowry, L., Mirkin, P. K., & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 21). January, 1980.

Deno, S. L., Marston, D., & Mirkin, P. K. Relationships among simple measures of written expression and performance on standardized achievement tests (Research Report No. 22). January, 1980.

Mirkin, P. K., Deno, S. L., Tindal, G., & Kuehnle, K. Formative evaluation: Continued development of data utilization systems (Research Report No. 23). January, 1980.

Deno, S. L., Robinson, S., Evans, P., & Mirkin, P. K. Relationships among classroom observations of social adjustment and sociometric rating scales (Research Report No. 24). January, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Factors influential on the psychoeducational decisions reached by teams of educators (Research Report No. 25). February, 1980.

Ysseldyke, J. E., & Algozzine, B. Diagnostic decision making in individuals susceptible to biasing information presented in the referral case folder (Research Report No. 26). March, 1980.

Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.

Ysseldyke, J. E., Regan, R. R., & Schwartz, S. Z. The use of technically adequate tests in psychoeducational decision making (Research Report No. 28). April, 1980.

Richey, L., Potter, M., & Ysseldyke, J. Teachers' expectations for the siblings of learning disabled and non-learning disabled students: A pilot study (Research Report No. 29). May, 1980.